

**REMARKS**

Reconsideration and allowance in view of the foregoing amendment and the following remarks are respectfully requested.

**Rejection of Claims 22-25, 27, 29-32 and 34 Under 35 U.S.C. §103(a)**

The Office Action rejects claims 22-25, 27, 29-32 and 34 under 35 U.S.C. §103(a) as being unpatentable over Ezzat et al. (Visual Speech Synthesis by Morphing Visemes) (“Ezzat et al.”) in view of Jiang et al. (Visual Speech Analysis with Application to Mandarin Speech Training) (“Jiang et al.”) in view of Hon et al. (Automatic Generation of Synthesis Units for Trainable Text-to-Speech Systems) (“Hon et al.”). Applicants thank the Examiner for the response to arguments found on pages 7 and 8 of the final Office Action. Applicants shall contact the Examiner and immediately request to schedule a personal interview with the Examiner and provide herein several comments to the Examiner’s response in order to set the stage for the interview.

The Examiner maintains that the Section 103 rejection that combines Ezzat et al. with Hon et al. is appropriate for two reasons. Applicants shall restate the reasons here and set forth our position and analysis regarding why Applicants believe Hon et al. should not be combined with Ezzat et al.

The Office Action first argues that because the prior art rejections presented in the Office Action do not explicitly rely upon Hon et al. for teaching diphones but rather on the unit selection process, the Office Action essentially argues that it is appropriate in a Section 103 analysis to ignore or discount portions of the reference that are not utilized or cited when rejecting the claims. Applicants strongly traverses this approach and notes that it is counter to numerous cases and citation within the MPEP to simply ignore portions of a reference in an analysis that involves determining whether it would be obvious to one of skill in the art to

combine two references. A simple example illustrates the slipper slope that this approach would cause. Assume an Office Action seeks to combine the Jones et al. patent with the Smith et al. patent. Further assume that Smith et al. expressly references problems within the Jones et al. approach. We can even assume that 90% of the discussion in Smith et al. specifically comments on why the Jones et al. approach is inferior and substandard. Then if we assume that buried within the Smith reference there is a paragraph that may teach a limitation in the claims being examined. Under the current Office Action's approach, all of the express teachings away from combining Smith et al. and Jones et al. can be simply ignored because the particular nugget of disclosure buried within Smith et al. is useful. However, as has been noted above, this cuts directly against the requirements of the MPEP to consider the prior art references in their entirety. Notably, MPEP 2141.02, Section VI., teaches that "a prior art reference must be considered in its entirety, i.e., as a whole, including portions that would lead away from the claimed invention." (emphasis in original) This plain statement in and of itself proves that the current analysis that ignores the express teachings in Hon et al. in which using diphones is discouraged cannot be maintained. Applicants respectfully submit that the Examiner must accept disclosure which is helpful in rejecting the claims but also must accept, analyze and process the disclosure that supports Applicants' position. Under such an analysis, the side with the weighted arguments by a preponderance of the evidence should prevail.

An even stronger citation is found in MPEP 2143.01, Section II. Here, the MPEP requires "the test for obvious is what the combined teachings of the references would have suggested to one of ordinary skill in the art, and all teachings in the prior art must be considered to the extent that they are in analogous arts. Where the teachings of two or more prior art references conflict, the Examiner must weight the power of each reference to suggest solutions to one of ordinary skill in the art, considering the degree to which one reference might accurately

discredit another.” (emphasis added) Applicants respectfully submit that this is a direct statement that applies to the current analysis in which the Examiner seeks to combine Hon et al. with Ezzat et al. Applicants have set forth teachings within the prior art that “conflict”. The Examiner must weight the suggestive power of each of these references to suggest solutions to one of ordinary skill in the art. In other words, the Examiner cannot simply ignore the teachings which would lead one of skill in the art away from combining these references.

Applicants further notes that the approach set forth in the MPEP more closely aligns with what would happen in the real world when a person of ordinary skill in the art might be reviewing or reading these papers. In such a case, the reader would certainly note in the initial section of Hon et al. the discouraging comments regarding the use of diphones. In that case, the mind of one of skill in the art would certainly have the suggestion in his or her mind against combining these references. For example, assume one of skill in the art was aware of the Ezzat et al. reference and the audio visual synchronization approach discussed in our earlier responses in which they highlight the good quality of using diphones when synchronizing between the audio and video. Then, if that person of skill in the art were to start to read Hon et al., they would immediately, even beginning on the first page, be met with criticisms of the diphone approach. The analysis that Applicants have in detail performed is the exact analysis performed by the MPEP in which the suggestive power of each reference has been analyzed and especially in light of how one reference discredits another. Accordingly, Applicants would respectfully request that rather than ignoring the teachings away from the combination of the references, that the Examiner either analyze them and explain why the preponderance of the evidence is in the Examiner’s favor or issue a Notice of Allowance. In sum, Applicants again reiterate that the Examiner cannot simply ignore the express teachings away from the combination of these references.

The Office Action next asserts that according to Section 2.2 of Hon et al., “the technology used in place of diphone enhances the use of diphones and is also well known in the art.” Applicants respectfully submit that this is an incorrect statement regarding the characterization and teachings of Hon et al. Applicants assert that the technology that is used in place of diphones is not taught as “enhancing the use of diphones”. Section 2.2 of Hon et al. teaches the following:

“To achieve a more natural voice quality, one must take more contexts into account, going beyond diphones. However, simply modeling *triphones* [9] (a phone with the immediate left and right contexts) already requires more than 10,000 units for English. Fortunately, effective clustering of similar contexts modeled in a sub-phonetic level, to allow flexible memory-quality compromise, has been well studied in the speech recognition community [6]. A senone [6][3] is a context-dependent sub-phonetic unit which is equivalent to a HMM state in a triphone. Senones can usually model spectral variation in a fine sub-phonetic level. However, they introduce many more junctions than a diphone system. Those excessive junctions create potential quality degradation for synthesis systems.”

Rather than somehow teaching that the Hon et al. technology which is used in place of diphones somehow “enhances” the use of diphones, Section 2.2 merely states that the suggestion of going beyond diphones and simply modeling triphones introduces many more junctions than a diphone system and thus can also potentially degrade the quality of a synthesis system. Applicants simply find no suggestion within Section 2.2 of Hon et al. technology that would enhance the use of diphones. Rather, they teach replacing the diphone approach as is set forth in Section 2.3 of Hon et al. rather than enhancing it. Accordingly, Applicants submit that this secondary factor cited in the Office Action to support maintaining the Section 103 reference is technically inaccurate and thus, must fall.

Next, under the second reason to maintain the rejection, the Office Action asserts that the Hon et al. technology is “also well known in the art.” Applicants are unsure what is really meant here inasmuch as the Office Action appears to seek to assert that “the technology used in place of diphones [that] enhances the use of diphones....is...well known in the art.” Since Applicants

have shown that Hon et al. fails to teach any such technology that “enhances” the use of diphones, Applicants also submit that such non-existent technology is not “well known in the art.” Thus, the other prong of this second reason to maintain the rejection also falls because the Examiner cannot rely on such non-existent technology being well known in the art to continue to support the conclusion that one of ordinary skill in the art can combine the references of Ezzat et al. and Hon et al. into a workable combination.

Applicants also note that MPEP 2143.01, Section III, expressly states that the mere fact that references “can” be combined does not render the resultant combination obvious unless the prior art also suggests the desirability of the combination. Again, this is highly applicable to the present analysis where the Office Action asserts that one of ordinary skill in the art “can combine the references” but Applicants have provided multiple and detailed reasons why the prior art fails to suggest the desirability of the combination.

The Office Action also includes the comment that enhancing qualities achieved beyond diphones would even motivate one of ordinary skill in the art to build the combination in order to achieve the higher quality output. Applicants submit that while as a general statement that may be true, when the suggestive power of Hon et al. and Ezzat et al. are analyzed in this context, they continue to teach away from their combination. For example, on page 46, right column, fourth paragraph, a triphone segment is referenced in Ezzat et al. Clearly, the ability of using triphones or units other than diphones were available and known by Ezzat et al. However, in their particular approach which requires audio visual synchronization as discussed in Section 7, they highlight the benefit of using diphones is important for the use of TTS timing and phonemic information in the manner of their research. The reason for concatenating diphones is because of the very good quality lip synchronization between the audio and the video. Thus, while Hon et al. teach a TTS engine that includes improvements over a diphone approach, Applicants maintain

that the requirement of Ezzat et al. differs from mere TTS but also requires synchronization with video. Clearly, since Ezzat et al. were aware other approaches besides using diphones, Applicants respectfully submit that for their particular requirement of using TTS timing and phonemic information for very good lip synchronization between the audio and video, that one of skill in the art would not be *motivated* or have *sufficient suggestion* to abandon the proven diphone approach of Ezzat et al. for the different approach highlighted in Hon et al. Accordingly, when the whole teachings of the prior art are properly analyzed for their suggestive power, Applicants respond to the Examiner's final point that one of skill in the art would not have the suggestion to abandon the diphone approach of Ezzat et al. for the teachings of Hon et al.

Applicants also highlight MPEP 2144.08, Section II.b. This portion of the MPEP requires the Office personnel to consider all rebuttal arguments and evidence presented by Applicants. Several portions of this reference highlight the fact that in consideration of rebuttal of evidence and arguments, Office personnel should weight the proper evidence and arguments and avoid giving the evidence no weight. Section III of this MPEP section also states that Office personnel should reconsider any initial obviousness determination in view of the entire record and all proposed rejection and their basis should be reviewed to confirm their correctness. Further, explicit finding on the motivation or suggestion to select the claimed invention should also be articulated in order to support a 35 U.S.C. ground of rejection. Applicants respectfully again notes that when the rebuttal arguments are considered and weighed in the balance under an appropriate preponderance of the evidence standard, Applicants submit that the weight of your evidence prevents the combination of Ezzat et al. and Hon et al. For these many reasons, Applicants respectfully submit that the present claims are patentable and in condition for allowance.

**Rejection of Claims 28 and 35 Under 35 U.S.C. §103(a)**

The Office Action rejects claims 28 and 35 under 35 U.S.C. §103(a) as being unpatentable over Ezzat et al. in view of Jiang et al. in further view of Hon et al. in further view of Brand (Voice Puppetry) ("Brand"). Applicants respectfully submit that based on the discussion above, that claims 28 and 35 are patentable inasmuch as these references should not be combined.

**CONCLUSION**

Having addressed all rejections and objections, Applicants respectfully submit that the subject application is in condition for allowance and a Notice to that effect is earnestly solicited. If necessary, the Commissioner for Patents is authorized to charge or credit the **Novak, Druce & Quigg, LLP, Account No. 14-1437** for any deficiency or overpayment.

Respectfully submitted,

Date: August <sup>21</sup>~~28~~, 2007

By: 

Correspondence Address:

Thomas A. Restaino  
Reg. No. 33,444  
AT&T Corp.  
Room 2A-207  
One AT&T Way  
Bedminster, NJ 07921

Thomas M. Isaacson

Attorney for Applicants  
Reg. No. 44,166  
Phone: 410-286-9405  
Fax No.: 410-510-1433